

⑫ 公開特許公報(A)

昭63-204300

⑤ Int. Cl.⁴G 10 L 9/18
H 04 B 1/66
14/04

識別記号

庁内整理番号

A-8622-5D
6745-5K
C-8732-5K

⑬ 公開 昭和63年(1988)8月23日

審査請求 有 発明の数 1 (全7頁)

⑥ 発明の名称 符号化伝送装置

⑦ 特 願 昭62-37304

⑧ 出 願 昭62(1987)2月20日

⑨ 発 明 者 谷 口 智 彦 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内
⑩ 発 明 者 伊 勢 田 衡 平 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内
⑪ 発 明 者 岡 崎 晃 二 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内
⑫ 発 明 者 天 野 文 雄 神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内
⑬ 出 願 人 富士通株式会社 神奈川県川崎市中原区上小田中1015番地
⑭ 代 理 人 弁理士 玉蟲 久五郎 外1名
最終頁に続く

明 細 書

1. 発明の名称 符号化伝送装置

2. 特許請求の範囲

音声信号を帯域圧縮して符号化して伝送し、復号信号をもとの帯域に伸張して再生する音声信号の符号化装置において、

音声信号の有声音の区間と無声音の区間とを識別する音声識別部(1)と、

有声音の区間において音声信号のピッチ周期を抽出するピッチ周期抽出部(2)と、

複数の該ピッチ周期分の音声信号から1ピッチ周期分をとり出して伝送する時間軸圧縮部(3)と、

無声音の区間において複数サンプル周期分の音声信号から1サンプル周期分をとり出して伝送する間引き部(4)とを送信側に具えとともに、

該1ピッチ周期分の信号からもとの複数ピッチ周期分に伸張する時間軸伸張部(5)と、

該1サンプル周期分の信号からもとの複数サン

プル周期分に伸張する補間部(6)とを受信側に具えてなることを特徴とする符号化伝送装置。

3. 発明の詳細な説明

(概 要)

音声信号の有声音の区間と無声音の区間とを識別し、音声信号の有声音の区間においては、ピッチ周期を利用して音声信号の時間軸方向の圧縮と伸張を行い、無声音の区間においては、サンプリング周期によつて音声信号の時間軸方向の圧縮と伸張を行うことによつて、有声音と無声音の両方の区間において、信号の時間軸方向の圧縮と伸張とが正しく行われるようにして、再生音全体の明瞭度を向上させる。

(産業上の利用分野)

本発明は音声信号の符号化伝送装置に係り、特に音声信号のピッチ周期における繰り返しを利用して時間軸方向に圧縮、伸張を行うTDHS(Time Domain Harmonic Scaling)方式高能率音声符

号化伝送装置に関するものである。

T D H S方式は、音声波形のピッチ周期による周期性を利用して時間軸方向に帯域の圧縮、伸張を行う符号化方式であつて、下記の文献(1)、(2)に詳細に記載されており、また文献(3)にその概要が記載されている。また文献(4)は同様に時間軸方向の圧縮、伸張を行う方式でT D H S方式の改良案の一つと言えるものである。

高能率音声符号化伝送装置はこのような方式を用い、音声信号をその品質を保つたまま情報圧縮を行う装置である。

移動通信、衛星通信あるいは企業内通信等における音声情報の伝送においては、回線コストの削減が重要であり、音声蓄積、音声応答システム等において音声情報の蓄積を行う場合には、蓄積用メモリ容量の削減が重要であるが、高能率音声符号化伝送装置はこれらの要求を可能にするものである。

しかしながらT D H Sを用いた高能率音声符号化伝送装置においては、無声音に対して再生音声

の明瞭度が悪いという問題があるが、このような音声の品質低下を生じないような音声信号の符号化伝送装置が要望される。

(参考文献)

- (1) D. Malah, "Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals" IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-27, pp. 121-133 Apr. 1979
- (2) R. V. Cox et al., "An Implementation of Time Domain Harmonic Scaling with Application to Speech Coding" ICC 82, pp. 4G.1.1-4
- (3) 古井貞熙「ディジタル音声処理」p. 122-124 東海大学出版会
- (4) 森田、板倉「自己相関法による音声の時間軸での伸縮方式とその評価」電気音響研究会資料 EA86-5

(従来の技術)

第5図は、従来のT D H S音声符号化伝送装置

3

4

BEST AVAILABLE COPY

の構成を示したものであつて、1は時間軸圧縮部(T D H C)、2はピッチ周期抽出部、3は符号化部、4は伝送路、5は復号化部、6は時間軸伸張部(T D H E)である。

第5図において、音声信号入力 $S(n)$ はピッチ周期抽出部2において、そのピッチ周期を抽出される。時間軸圧縮部1は抽出されたピッチ周期に応じて入力 $S(n)$ を時間軸方向に圧縮し、圧縮信号 $Sc(n)$ を生じる。圧縮信号 $Sc(n)$ は符号化部3に加えられて、任意の符号化方式で符号化される。

符号化された信号は、伝送路4を経て伝送される。受信側においては、復号化部5においてこれを復号して圧縮再生信号 $\hat{Sc}(n)$ を生じる。この信号は時間軸伸張部6に加えられて別に伝送されたピッチ周期の信号を用いて時間軸方向に伸張されて、再生出力 $\hat{S}(n)$ を生じる。

第6図は第5図の符号化伝送装置における時間軸圧縮と時間軸伸張の処理の例を説明するものであつて、(a)は時間軸圧縮を示し、(b)は時間軸伸張を示している。

5

すなわち時間軸圧縮を行う場合は、第6図(a)のように抽出されたピッチ周期 P ごとに入力 $S(n)$ を切り出し、2周期を1組として前の周期の信号には重み窓 $W(n)$ をかけ、後の周期の信号には逆特性の重み窓 $1-W(n)$ をかけた出力をそれぞれ発生し、これを平均化して1周期の信号を得ることによつて、 $1/2$ に圧縮が行われた圧縮信号 $Sc(n)$ が得られる。

一方時間軸伸張を行う場合は、第6図(b)のように圧縮再生信号 $\hat{Sc}(n)$ を3周期分とり、前の2周期に重み窓 $1-W(n)$ をかけ、後の2周期に重み窓 $W(n)$ をかけた出力をそれぞれ発生し、これを平均化することによつて2周期の再生信号 $\hat{S}(n)$ を得ることによつて、もとの帯域に伸張される。

(発明が解決しようとする問題点)

従来のT D H S方式においては、音声信号の周期性を抽出するための評価関数として、(1)式のように波形の相関を算出するもの、または(2)式のように波形の類似性を算出するものを用い、これに

6

よつて最も相関が大きい周期をピッチ周期と
していた。

$$S_1(N) = \sum X_j \cdot X_{j-N} / \sum X_j^2 \quad \dots (1)$$

$$S_2(N) = \sum |X_j - X_{j-N}| \quad \dots (2)$$

またこのようなピッチ周期の探索の範囲として
は、ピッチ周波数の上限、下限を規定し、例えば
 $16 \leq N \leq 200$ 程度の範囲で探索を行っていた。

このような探索手法では、周期性のない無声音
のような場合、正しく周期の抽出を行うことがで
きず、従つてランダムな値に設定されたピッチ周
期に応じて圧縮、伸張が行われるため、波形の性
質が保存されず再生音声の明瞭度が悪くなる等の
品質劣化を生じていた。

(問題点を解決するための手段)

本発明はこのような従来技術の問題点を解決し
ようとするものであつて、第1図に示す原理的構
成を有し、音声信号を帯域圧縮して符号化して伝
送し、復号信号をもとの帯域に伸張して再生する
音声信号の符号化装置において、音声識別部10と、

ピッチ周期抽出部2と、時間軸圧縮部1と、間引
き部7とを送信側に具えるとともに、時間軸伸張
部6と、補間部8とを受信側に具えたものである。

音声識別部10は、音声信号の有声音の区間と無
声音の区間とを識別する。

ピッチ周期抽出部2は、有声音の区間において
音声信号のピッチ周期を抽出する。

時間軸圧縮部1は、複数ピッチ周期分の音声信
号から1ピッチ周期分をとり出して伝送する。

間引き部7は、無声音の区間において複数サン
プル周期分の音声信号から1サンプル周期分をと
り出して伝送する。

時間軸伸張部6は、1ピッチ周期分の信号から
もとの複数ピッチ周期分に伸張する。

補間部8は、1サンプル周期分の信号からもと
の複数サンプル周期分に伸張する。

(作 用)

音声信号の有声音の区間と無声音の区間とを識
別し、音声信号の有声音の区間においては、送信

7

8

BEST AVAILABLE COPY

側でピッチ周期を抽出して複数ピッチ周期分の音
声信号から1ピッチ周期分をとり出すことによつ
て時間軸方向の圧縮を行つて送出し、受信側で伝
送された1ピッチ周期分の信号からもとの複数ピ
ッチ周期分に時間軸方向に伸張する。

一方無声音の区間においては、送信側で複数サ
ンプル周期分の音声信号から1サンプル周期分を
とり出すことによつて時間軸方向の圧縮を行つて
送出し、受信側で伝送された1サンプル周期分の
信号からもとの複数サンプル周期分に時間軸方向
に伸張する。

これによつて有声音と無声音の両方の区間にお
いて、信号の時間軸方向の圧縮と伸張とが正しく
行われるので、再生音全体の明瞭度が向上する。

(実施例)

第2図は本発明の一実施例の構成を示したもの
であつて、第5図における同じ部分を同じ番号
で示し、7は間引き部、8は補間部、9₁、9₂
はスイッチである。

第3図は第2図の実施例における間引き部7、
補間部8の処理を説明する図である。

また第4図は本発明によるピッチ周期抽出部2
の構成例を示したものであつて、11は共分散算出
部、12は最大値検出部、13は共分散閾値設定部、
14は比較部、15は判定部、16₁、16₂はスイッチ
である。

第2図において、ピッチ周期抽出部2は入力S
(n)の周期性を検出して、周期性が大きい有声音の
部分に対してはそのピッチ周期P=p(pはピッ
チ周期に対応するサンプリング周期数)を抽出し
て出力し、周期性が小さい有声音の部分に対して
はピッチ周期P=1を出力する。

スイッチ9₁、9₂はピッチ周期P≠1のとき
は、それぞれ時間軸圧縮部1、時間軸伸張部6の
側に切り替えられ、時間軸圧縮部1はピッチ周期
P=pによつて第5図の場合と同様にして時間軸
圧縮を行い、時間軸伸張部6は同様にピッチ周期
P=pによつて時間軸伸張の処理を行う。

一方ピッチ周期P=1のときは、スイッチ9₁、

9

10

9₂はそれぞれ間引き部と補間部8の側に切り替えられ、ピッチ周期 $P=1$ によつて間引きと補間の処理を行う。ここでピッチ周期 $P=1$ は音声のサンプリング周期によつて処理を行うことを意味する。間引き処理においては、(3)式に示すように2サンプリング周期の音声信号の平均値によつて、1サンプリング周期の圧縮信号 $Sc(n)$ を発生することによつて、1/2に圧縮された圧縮信号 $Sc(n)$ を得る。

$$Sc(i) = (S(i) + S(i+1)) / 2 \quad \dots (3)$$

第3図(a)は、このような処理を模式的に示している。

補間部8においては、圧縮再生信号 $\hat{Sc}(i)$ に対し1サンプリング周期後の信号を用いて(4)₁、(4)₂式に従つて補間処理を行つて、1サンプリング周期の圧縮再生信号 $\hat{Sc}(n)$ から2サンプリング周期の再生出力 $S(n)$ を発生することによつて、もとの帯域に伸張する。

$$\hat{Sc}(2i-1) = (\hat{Sc}(i-1) + 3 \cdot \hat{Sc}(i)) / 4 \quad \dots (4)_1$$

$$\hat{Sc}(2i) = (3 \cdot \hat{Sc}(i) + \hat{Sc}(i+1)) / 4 \quad \dots (4)_2$$

第3図(b)は、このような処理を模式的に示している。

またこの場合におけるピッチ周期の抽出は、第4図に示すようにして行われる。すなわち共分散算出部11は入力 $S(n)$ に対して、(5)式で定義される共分散を M_1 次から M_2 次まで算出する。

$$C(n) = \sum S(i) \cdot S(i+n) / \sqrt{\sum S(i)^2 \cdot \sum S(i+n)^2} \quad \dots (5)$$

ここで通常 $M_1 = 16$ 、 $M_2 = 200$ 程度である。最大値検出部12は、このようにして求められた $C(M_1) \sim C(M_2)$ のうちの最大値 $\max C(i)$ ($i = M_1 \sim M_2$)を検出して、共分散値 $C(P)$ とする。ここで P はピッチ周期である。

このようにして検出された共分散値 $C(P)$ を比較部14に加え、共分散値閾値設定部13に設定された一定の閾値 C_{th} と比較する。比較部14は $C(P) < C_{th}$ または $C(P) \geq C_{th}$ に応じて出力を発生する。ここで $C(P) < C_{th}$ ならば周期性が小さいので無声音であり、 $C(P) \geq C_{th}$ ならば周期性が大きいのので有声音であるから、判定部15は比較部14の出力に応じて無声音、有声音の判定を行つて、有声音の場合

1 1
BEST AVAILABLE COPY
1 2

合にピッチ周期 $P=p$ (p はサンプリング周期数で示されたピッチ周期)を出力し、無声音の場合は $P=1$ を出力する。この場合、共分散値閾値設定部13に設定される閾値 C_{th} の値は、通常0.6~0.7程度である。

〔発明の効果〕

以上説明したように本発明によれば、従来、時間軸圧縮、伸張方式において問題となつていた、周期性のない無声音の区間に対する処理を、この区間を検出して間引き、補間の処理に置き換えることによつて、無声音に対する再生音声品質が改善され、再生音全体の明瞭度を向上させることができる。

4.図面の簡単な説明

第1図は本発明の原理的構成を示す図、

第2図は本発明の一実施例を示す図、

第3図は間引き部と補間部の処理を説明する図、

第4図はピッチ周期抽出部の構成例を示す図、

第5図は従来の符号化伝送装置の構成を示す図、

第6図は第5図の符号化伝送装置における時間軸圧縮と時間軸伸張の処理の例を説明する図である。

1…時間軸圧縮部(TDHC)

2…ピッチ周期抽出部

3…符号化部

4…伝送路

5…復号化部

6…時間軸伸張部(TDHE)

7…間引き部

8…補間部

9₁、9₂…スイッチ

11…共分散算出部

12…最大値検出部

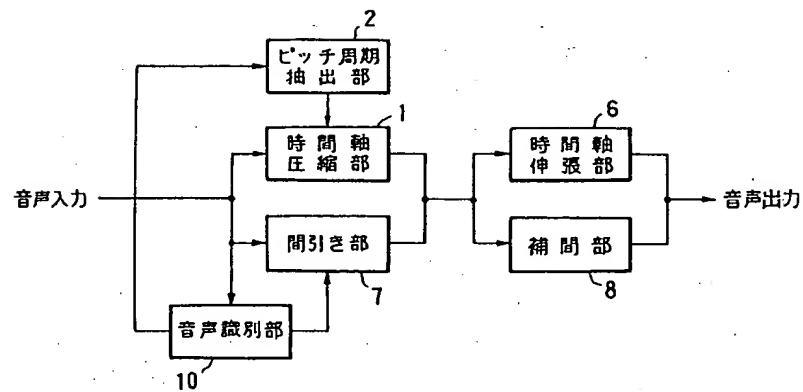
13…共分散閾値設定部

14…比較部

15…判定部

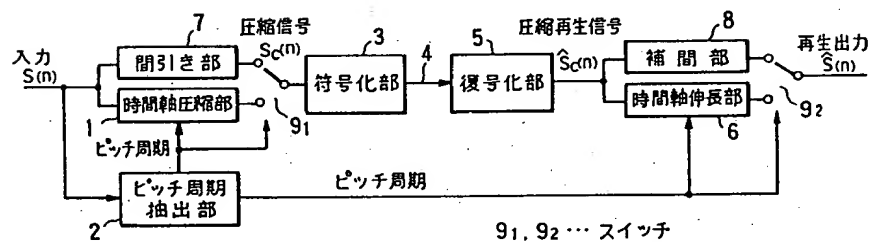
16₁、16₂…スイッチ

BEST AVAILABLE COPY



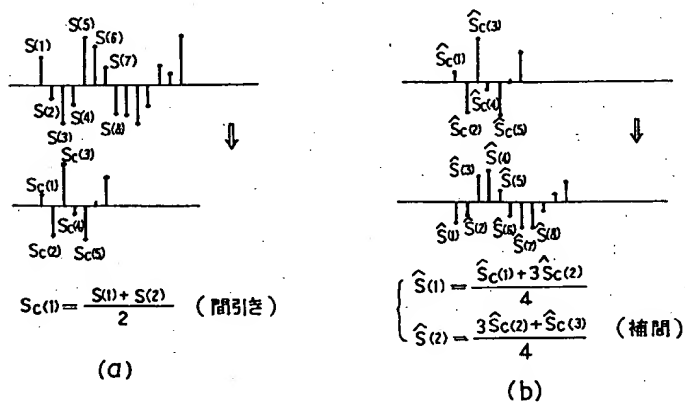
本発明の原理的構成を示す図

第 1 図



本発明の一実施例を示す図

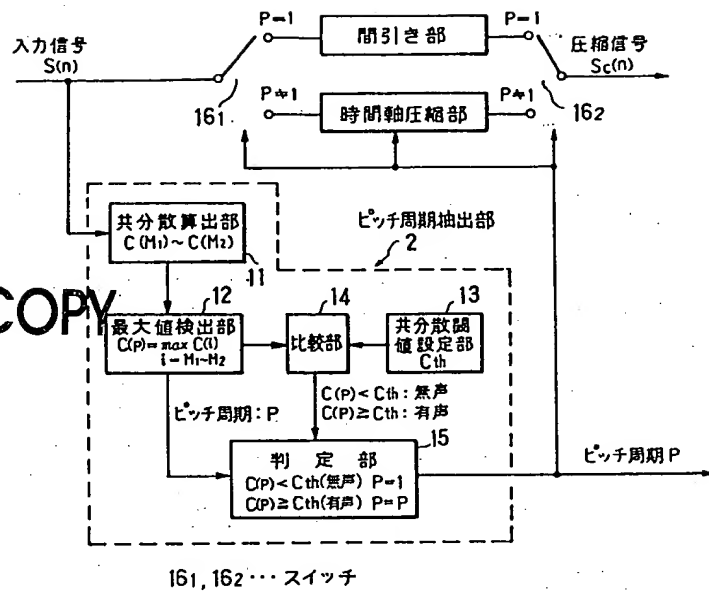
第 2 図



間引き部と補間部の処理を説明する図

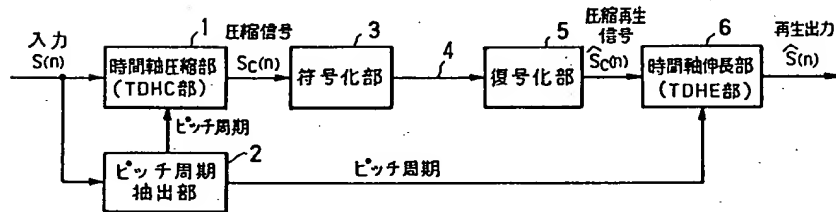
第 3 図

BEST AVAILABLE COPY



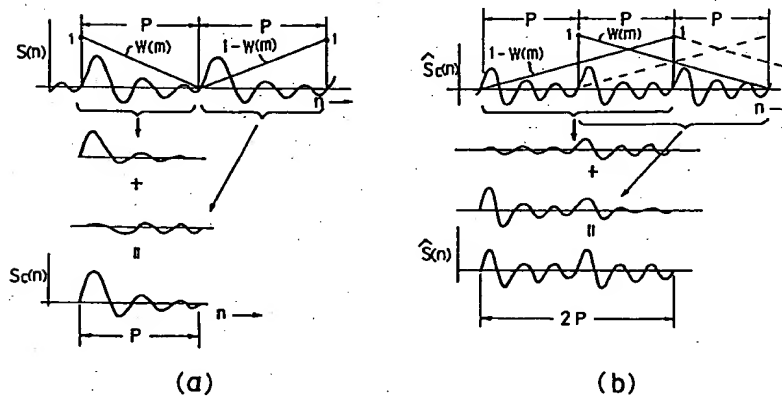
ピッチ周期抽出部の構成例を示す図

第 4 図



従来の符号化伝送装置を示す図

第 5 図



第5図の符号化伝送装置における時間軸圧縮と時間軸伸長の処理を説明する図

第 6 図

第1頁の続き

②発 明 者

海 上

重 之

神奈川県川崎市中原区上小田中1015番地 富士通株式会社
内

BEST AVAILABLE COPY